

# 空氣品質指標 (AQI)

資料來源：政府資料開放平臺

# 大綱

- ❑ 研究動機
- ❑ 資料來源
- ❑ 工作分配
- ❑ 方法簡介
- ❑ 初步結果

# 研究動機

## 什麼是「空氣品質指標（AQI）」？

定量描述空氣品質狀況的非線性無因次指數。其數值越大、等級和類別越高、顏色越深，代表空氣污染狀況越嚴重，對人體的健康危害也就越大。 -維基百科

- 風速與 AQI 值的關係？
- 是否可以利用現有的數據猜測污染物？

# 資料來源

## 空氣品質指標(AQI)

每小時提供各測站之空氣品質指標 (AQI) , 原始資料版本公告於空氣品質監測網<https://airtw.moeenv.gov.tw>

評分此資料集：



平均 3.17 (283 人次投票)



瀏覽次數：325632

下載次數：83753

意見數：55

氣候環境

DCAT 詞彙

列印

主要欄位說明  
\*粗體欄位為資料  
標準欄位

sitename(測站名稱)、county(縣市)、aqi(空氣品質指標)、pollutant(空氣污染指標物)、status(狀態)、**so2(二氧化硫 (ppb))**、**co(一氧化碳 (ppm))**、**o3(臭氧 (ppb))**、**o3\_8hr(臭氧8小時移動平均 (ppb))**、**pm10(懸浮微粒 (μg/m3))**、**pm2.5(細懸浮微粒 (μg/m3))**、**no2(二氧化氮 (ppb))**、**nox(氮氧化物 (ppb))**、**no(一氧化氮 (ppb))**、wind\_speed(風速 (m/sec))、wind\_dir(風向 (degrees))、publishtime(資料發布時間)、co\_8hr(一氧化碳8小時移動平均 (ppm))、pm2.5\_avg(細懸浮微粒移動平均值 (μg/m3))、pm10\_avg(懸浮微粒移動平均值 (μg/m3))、so2\_avg(二氧化硫移動平均值 (ppb))、longitude(經度)、latitude(緯度)、siteid(測站編號)

資料資源下載網  
址



檢視資料

空氣品質指標(AQI)-CSV



檢視資料

空氣品質指標(AQI)-JSON



檢視資料

空氣品質指標(AQI)-XML

提供機關

環境部

提供機關聯絡人  
姓名

敖先生 (02-2311-7722#2345)

更新頻率

每1時

授權方式

政府資料開放授權條款-第1版

計費方式

免費

連接：<https://data.gov.tw/dataset/40448>

# 工作分配

李宇捷  
01057108

資料收集

研究模型

簡報製作

陳柏瑋  
01057138

資料處理

研究模型

簡報製作

洪旻昌  
01057158

特徵分析

研究模型

簡報製作

# 方法簡介

## 資料集欄位介紹

欄位名稱	欄位描述
sitename	測站名稱
county	縣市
aqi	空氣品質指標
pollutant	空氣污染指標物
status	狀態
so2	二氧化硫 (ppb)
co	二氧化碳 (ppm)
o3	臭氧 (ppb)
o3_8hr	臭氧8小時移動平均 (ppb)
pm10	懸浮微粒 ( $\mu\text{g}/\text{m}^3$ )
pm2.5	細懸浮微粒 ( $\mu\text{g}/\text{m}^3$ )

欄位名稱	欄位描述
no2	二氧化氮 (ppb)
nox	氮氧化物 (ppb)
no	一氧化氮 (ppb)
wind_speed	風速 (m/sec)
wind_dirac	風向 (degrees)
publishtime	資料發佈時間
co_8hr	一氧化碳8小時移動平均 (ppm)
pm2.5_avg	細懸浮微粒移動平均值 ( $\mu\text{g}/\text{m}^3$ )
pm10_avg	懸浮微粒移動平均值 ( $\mu\text{g}/\text{m}^3$ )
so2_avg	二氧化碳移動平均值 (ppb)

欄位名稱	欄位描述
longitude	經度
latitude	緯度
siteid	測站編號

# 方法簡介

## 收集資料

時間範圍: 2025/03/24 00:00 至 2025/03/30 23:00

測站 : 87 站 (22 縣市)

總筆數據: 14,617 筆 (每站平均: 168 筆數據)

## 資料屬性

sitename	{ 22 測站名稱 }			
county	{ 22 縣市名稱 }			
pollutant	{ 細懸浮微粒、懸浮微粒、二氧化氮、臭氣八小時、二氧化硫、臭氧 }			
status	{ 良好、普通、對敏感族群不健康、對所有族群不健康 }			
datacreationdate	{ 2025/03/24 00:00 ~ 2025/03/30 23:00 }			
aqi	{ -1 ~ 183 }	no2	{ 0 ~ 72 }	pm10_avg { 0 ~ 122.9 }
so2	{ 0 ~ 68.9 }	nox	{ 0 ~ 172.8 }	so2_avg { 0 ~ 10 }
co	{ 0.06 ~ 2.99 }	no	{ 0 ~ 122.9 }	longitude { 118.31 ~ 121.82 }
o3	{ 0 ~ 131 }	windspeed	{ 0 ~ 588 }	latitude { 21.95 ~ 26.15 }
o3_8hr	{ 0 ~ 95 }	winddirect	{ 0 ~ 67 }	siteid { 1 ~ 313 }
pm10	{ 0 ~ 588 }	co_8hr	{ 0 ~ 72 }	
pm2.5	{ 0 ~ 67 }	pm2.5_avg	{ 0 ~ 172.8 }	

# 方法簡介

## 資料處理

- 欄位篩選

研究問題	輸入（屬性）	輸出
風速與 AQI 值 的關係	wind_speed aqi	pattern
猜測污染物	so2 co o3 pm10 pm2.5 no2 nox no	pollutant（屬性）

- 欄位缺值 → pollutant 補值(當作一個類別)、其他手動刪掉整列



# 方法簡介

## 風速與 AQI 值

Classification : Linear Regression

Select Attribute : Correlation Attribute Evaluation

## 猜測污染物

Classification : NaiveBayes

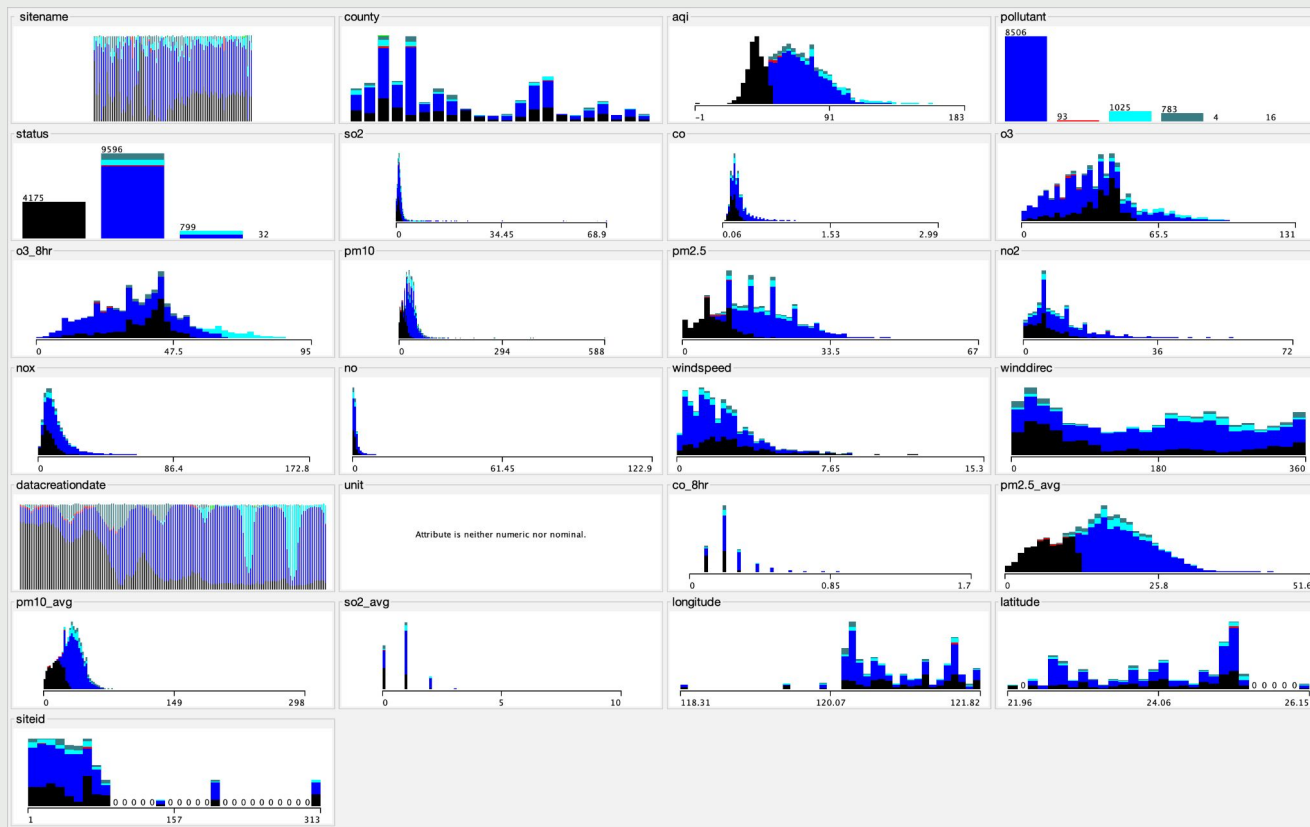
ZeroR

SMO

REPTree

J48

## 初步結果



# 初步結果

## 風速與 AQI 值

### Select Attribute : Correlation Attribute Evaluation

Attribute Evaluator

Choose **CorrelationAttributeEval**

Search Method

Choose **Ranker -T -1.7976931348623157E308 -N -1**

Attribute Selection Mode

☒ Use full training set

☐ Cross-validation

Folds **10**

Seed **1**

No class ☒

**Start** **Stop**

Result list (right-click for options)

21:44:08 - Ranker + CorrelationAttributeEval

Attribute selection output

```
=== Run information ===

Evaluator:   weka.attributeSelection.CorrelationAttributeEval
Search:     weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N -1
Relation:   20250324_20250330-FILTERED-1
Instances:  14224
Attributes: 2
            aqi
            windspeed
Evaluation mode: evaluate on all training data

=== Attribute Selection on all input data ===

Search Method:
  Attribute ranking.

Attribute Evaluator (supervised, Class (numeric): 2 windspeed):
  Correlation Ranking Filter

Ranked attributes:
-0.217 1 aqi

Selected attributes: 1 : 1
```

風速越小，AQI 值越大

# 初步結果

## 風速與 AQI 值

### Classification : Linear Regression

Classifier

Choose **LinearRegression** -S 0 -R 1.0E-8 -num-decimal-places 4

Test options

☐ Use training set

☐ Supplied test set

☒ Cross-validation Folds

☐ Percentage split %

(Num) aqi

Result list (right-click for options)

22:20:37 - functions.LinearRegression

Classifier output

=== Run information ===

Scheme: weka.classifiers.functions.LinearRegression -S 0 -R 1.0E-8 -num-decimal-places 4  
Relation: 20250324\_20250330-FILTERED-1  
Instances: 14224  
Attributes: 2  
aqi  
windspeed

Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

Linear Regression Model

aqi =

$-2.4818 * \text{windspeed} + 70.4164$

Time taken to build model: 0 seconds

=== Cross-validation ===

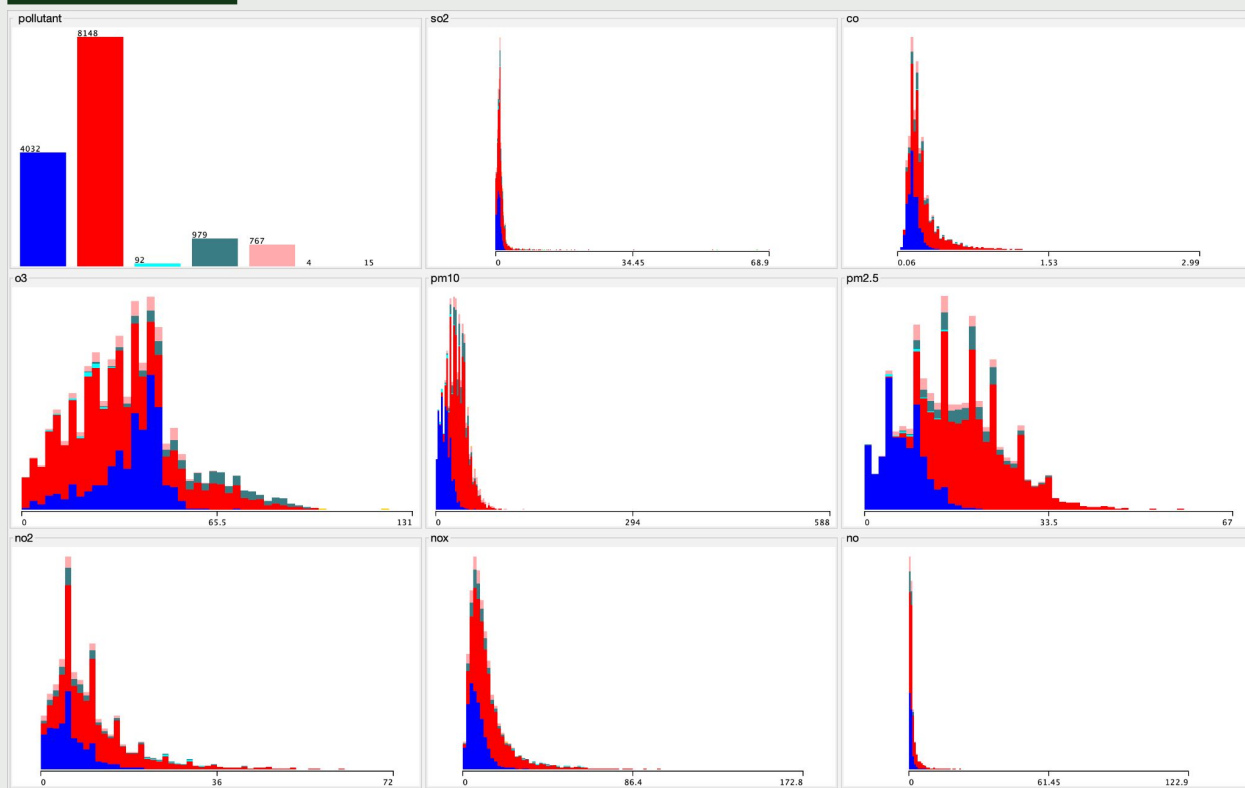
=== Summary ===

Correlation coefficient	0.2167
Mean absolute error	16.8507
Root mean squared error	21.1851
Relative absolute error	97.2511 %
Root relative squared error	97.6164 %
Total Number of Instances	14224

# 初步結果

## 猜測污染物

### 特徵分析



# 初步結果

猜測污染物

Classification : Comparison

Model	Accuracy (%)
NaiveBayes	57.512
ZeroR	58.047
SMO	79.041
<b>REPTree</b>	<b>80.53</b>
J48	80.067

# 初步結果

## 猜測污染物

### Classification : REPTree (10-Folds Cross-Validation)

Classifier output

Time taken to build model: 0.04 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	11304	80.53	%
Incorrectly Classified Instances	2733	19.47	%
Kappa statistic	0.6404		
Mean absolute error	0.0751		
Root mean squared error	0.2043		
Relative absolute error	45.9169	%	
Root relative squared error	71.4408	%	
Total Number of Instances	14037		

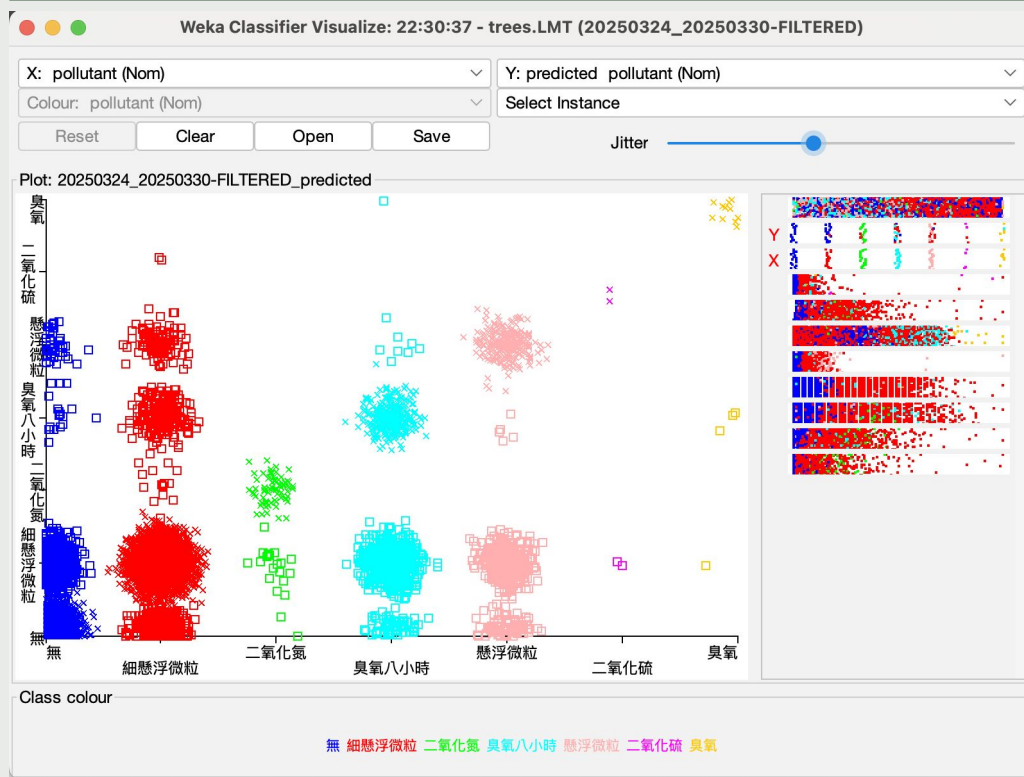
=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.821	0.057	0.852	0.821	0.836	0.772	0.954	0.894	無
	0.903	0.288	0.813	0.903	0.856	0.634	0.890	0.899	細懸浮微粒
	0.815	0.001	0.862	0.815	0.838	0.837	0.968	0.841	二氧化氮
	0.301	0.019	0.546	0.301	0.388	0.374	0.885	0.410	臭氧八小時
	0.329	0.015	0.556	0.329	0.413	0.403	0.843	0.373	懸浮微粒
	0.000	0.000	?	0.000	?	?	0.789	0.002	二氧化硫
	0.800	0.000	0.800	0.800	0.800	0.800	0.933	0.598	臭氧
Weighted Avg.	0.805	0.186	?	0.805	?	?	0.906	0.834	

# 猜測污染物

## Classification : REPTree (10-Folds Cross-Validation)

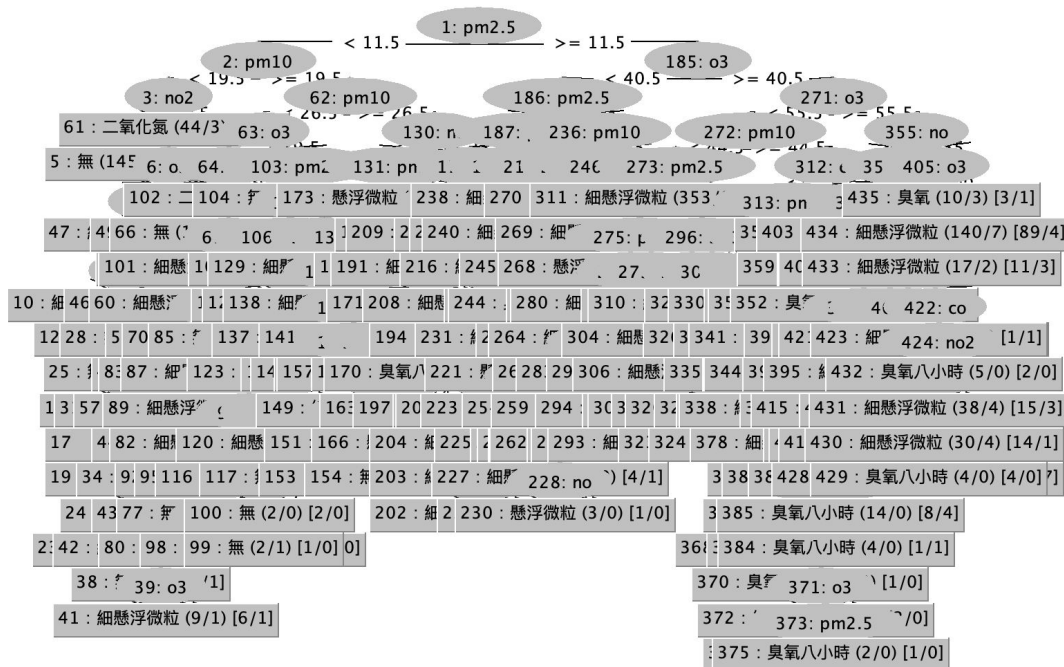
# 初步結果





# 初步結果

## Tree View



謝謝聆聽～